The model of classical communication complexity was defined by Yao in 1979 [5] and is still a very active area of research in computer science. In addition to many applications in communication protocols, this model also has some important applications in circuit complexity. In the basic model, we have two parties, Alice and Bob, who want to compute a function $f : X \times Y \to \{0,1\}$, where typically $X = Y = \{0,1\}^n$. Initially, Alice is given $x \in X$ and Bob is given $y \in Y$. They are then allowed to send bits to each other, and at the end of the protocol, Bob should output a guess for $f(x,y)$ which is correct with probability at least, say, $2/3$. We define the *communication complexity* of a function f as the minimum number of bits needed to compute it, and denote it by $R(f)$. It is important to remember that there are no computational restrictions on the parties, so we can think of them as being all-powerful.

Clearly, the communication complexity of any f is at most n, since there always exists the trivial protocol in which Alice sends her entire input to Bob, who can then compute $f(x,y)$ by himself. However, in many interesting cases, considerably less communication suffices.

In this lecture we will consider the model of *quantum* communication complexity. This model was introduced in 1993, again by Yao [6]. Here, we allow the parties to transfer qubits instead of bits, and as before, there are no computational limitations imposed on the parties. We define the quantum communication complexity of a function f as before, and denote it by $Q(f)$. Clearly, the quantum model is not weaker than the classic one, as any classical communication protocol with communication c is in particular a quantum communication protocol with communication c. In other words, we have $Q(f) \le R(f)$ for any function f. In fact, quantum protocols are sometimes highly superior to classical ones: there are certain (partial) functions f that require exponentially more communication in the classical model than in the quantum model.

There are several variants of the basic communication model described above. In one such variant, we allow the parties to share an arbitrarily long random bit-string. A potentially more powerful variant is obtained when we allow the parties to share an unlimited number of EPR pairs. Although these variants are very interesting, in this lecture we will mostly focus on the basic model where no shared randomness nor entanglement is allowed.

To get a feel for the model, let us say a few words on the equality function: $\mathsf{EQ}(x,y) = 1$ iff $x = y$. If one insists on a protocol with no error, the trivial protocol needs n bits of communication, and this is optimal. For the more interesting case where a small error probability is allowed, there are much better protocols: let Alice and Bob use their shared randomness to choose a random string $z \in \{0,1\}^n$. Alice then sends Bob the inner product $\langle x, z\rangle$. Bob compares the received bit to $\langle y, z\rangle$ and outputs 1 iff the bits are equal. It is easy to see that if $x = y$ Bob always outputs 1, whereas if $x \ne y$, Bob outputs 0 with probability $1/2$. One can amplify the probability of success by using, say, two strings z_1, z_2 instead of one. So only a constant number of bits is required if Alice and Bob have shared randomness. Moreover, by a theorem of Newman, one can avoid the use of shared randomness while increasing the communication by only $O(\log n)$. So we have that $R(\mathsf{EQ}) = O(\log n)$ (this result can also be derived directly, without using Newman's theorem, by using error correcting codes). Clearly, this also implies $Q(\mathsf{EQ}) = O(\log n)$. Both results are essentially tight.

Our focus in this lecture will be on another two important functions: the inner product function (IP) and the disjointness function (DISJ). The former is defined as

$$\mathsf{IP}(x,y) = 1 \Leftrightarrow \sum_{i=1}^{n} (x_i \wedge y_i) \equiv 1 \pmod 2,$$

i.e., it is one iff the number of places where both strings are 1 is odd. It is known that $R(\mathsf{IP}) = \Theta(n)$, hence randomized protocols cannot do much better than the trivial n. In this lecture we will show that even quantum protocols cannot help: $Q(\mathsf{IP}) = \Theta(n)$. The second function we consider is the disjointness function, defined as

$$\mathsf{DISJ}(x,y) = 1 \Leftrightarrow \forall i : x_i = 0 \vee y_i = 0.$$

If we think of x and y as subsets of $\{1, \ldots, n\}$, then the disjointness function indicates if the two sets are disjoint. Here, the situation is more interesting: A tight classical bound of $R(\mathsf{DISJ}) = \Theta(n)$ was shown in 1992 [3], so classical protocols cannot do much better than the trivial protocol. In the quantum model, Buhrman, Cleve, and Wigderson [2] showed how to convert Grover's algorithm into a communication protocol with communication $O(\sqrt{n}\log n)$. This has since been improved to $O(\sqrt{n})$ by Aaronson and Ambianis [1], matching the lower bound of $\Omega(\sqrt{n})$ found by Razborov [4]. So to conclude, we have that $Q(\mathsf{DISJ}) = \Theta(\sqrt{n})$ whereas $R(\mathsf{DISJ}) = \Theta(n)$. This quadratic gap between the classical and the quantum communication complexity is currently the best known gap for a *total* function (i.e., functions defined on the entire set $X \times Y$). Before going on, let us quickly recall the elegant protocol of Buhrman et al.

THEOREM 1 $Q(\mathsf{DISJ}) = O(\sqrt{n}\log n)$.

PROOF: Our goal is to design a communication protocol that allows Alice and Bob to decide if there exists an i such that $x_i = y_i = 1$, or equivalently, $x_i \wedge y_i = 1$. Recall that Grover's algorithm finds an index i such that $z_i = 1$ using $O(\sqrt{n})$ oracle calls, assuming such an i exists. In the following, we will show how to use Grover's algorithm to solve the DISJ problem.

The idea of the protocol is to let Alice simulate Grover's algorithm with the oracle given by $z_i = x_i \wedge y_i$. Recall that Grover's algorithm consists of a sequence of unitary operations, interleaved with oracle calls. Alice can perform the unitary operations herself. The thing that needs to be shown is how to implement an oracle call, i.e., the unitary transformation that acts by

$$|i, b\rangle \mapsto |i, b \oplus (x_i \wedge y_i)\rangle.$$

This is done in three steps. In the first step, Alice performs the mapping

$$|i, b\rangle \mapsto |i, b, x_i\rangle.$$

She then sends this register to Bob, who performs the mapping

$$|i, b, x_i\rangle \mapsto |i, b \oplus (x_i \wedge y_i), x_i\rangle.$$

Finally, he returns this register to Alice, who erases the third register by XORing x_i into it,

$$|i, b \oplus (x_i \wedge y_i), x_i\rangle \mapsto |i, b \oplus (x_i \wedge y_i)\rangle.$$

The total communication complexity is $O(\sqrt{n}\log n)$, since we have $O(\sqrt{n})$ oracle calls, each involving a communication of $O(\log n)$ qubits between Alice and Bob. □

1 The Model

We now give a more formal definition of the quantum communication complexity model. A quantum communication protocol is composed of a sequence of arbitrary unitary transformation $U_1, U_2, \ldots, U_q$ applied alternatively by Alice and Bob (see Figure 1). These transformations act on the Hilbert space $\mathcal{H}_A \otimes \mathbb{C}^2 \otimes \mathcal{H}_B$. The initial state is $|0, x\rangle_A \otimes |0\rangle \otimes |0, y\rangle_B$ corresponding to Alice's state, which contains a possibly large number of ancillas in addition to her input, the communication qubit and Bob's state. The amount of communication is defined as q. The output is written in the communication qubit in the last stage of the protocol. We remark that one can also define a model where the channel contains more than one qubit, but the two models are essentially equivalent.

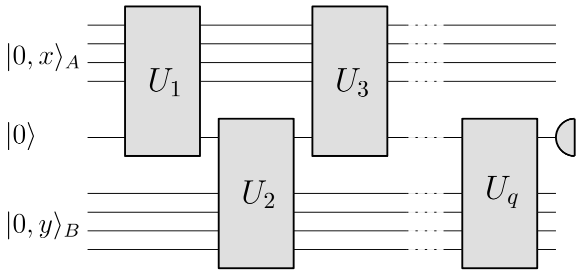

Figure 1: A communication protocol

2 Matrix Norms

We first introduce several mathematical tools that will help us afterwards in the analysis.

THEOREM 2 (SINGULAR VALUE DECOMPOSITION) *Any matrix A can be written as $A = UDV$ for unitary matrices U, V and a positive diagonal matrix D. The values on the diagonal of D are known as the* singular values *of A. Equivalently,*

$$A = \sum_{i=1}^{rank} s_i |\phi_i\rangle\langle\psi_i|$$

where $\{|\phi_i\rangle\}$ and $\{|\psi_i\rangle\}$ are sets of orthonormal vectors (corresponding to the columns of U and of V respectively), and $s_i > 0$ are the singular values of A.

Based on the vector of singular values, we can define some natural norms on matrices. We will be interested in three norms, corresponding to the ℓ_1, ℓ_2, and ℓ_∞ norms,

$$\begin{aligned} \|A\|_{\mathrm{tr}} &:= \sum s_i \\ \|A\|_F &:= \sqrt{\sum s_i^2} = (\mathrm{tr}(A^\dagger A))^{1/2} = \Big(\sum_{i,j} |a_{ij}|^2\Big)^{1/2} \\ \|A\| &:= \max s_i = \max_{x:\|x\|=1} \|Ax\|, \end{aligned}$$

where the identities follow easily from the singular value decomposition, and in the third line the vector norm is the ℓ_2 norm. These norms are known as the *trace norm*, the *Frobenius norm*, and the *operator norm*. An important property common to all three norms is that they are invariant under (both left and right) unitary operations, i.e., $|||UAV||| = |||A|||$ where $|||\cdot|||$ is any of these norms. This property follows immediately from the uniqueness (up to permutation) of the vector of singular values.

In addition, we define the inner product of two (not necessarily square) matrices of the same dimensions:

$$\langle A, B\rangle = \sum_{ij} a_{ij}^* b_{ij} = \mathrm{tr}(A^\dagger B),$$

which is also invariant under unitary transformations

$$\langle UAV, UBV\rangle = \mathrm{tr}(V^\dagger A^\dagger U^\dagger UBV) = \mathrm{tr}(A^\dagger B) = \langle A, B\rangle.$$

Notice that $\langle A, A\rangle = \|A\|_F^2$. We will need the following inequality (for diagonal matrices, this inequality is essentially an easy inequality on vectors).

LEMMA 3 $|\langle A,B\rangle| \le \|A\|_{\rm tr}\cdot\|B\|$

PROOF: Write $A=\sum s_i|\phi_i\rangle\langle\psi_i|$ in its singular value decomposition. Using the triangle inequality, we have

$$\begin{aligned}|\langle A,B\rangle| = |\mathrm{tr}(A^\dagger B)| &\le \sum s_i\big|\mathrm{tr}(|\phi_i\rangle\langle\psi_i|B)\big| = \sum s_i\big|\mathrm{tr}(\langle\psi_i|B|\phi_i\rangle)\big| \\ &= \sum s_i\big|\langle\psi_i|B|\phi_i\rangle\big| \le \sum s_i\big\|B|\phi_i\rangle\big\| \le \sum s_i\|B\| \\ &= \|A\|_{\rm tr}\cdot\|B\|.\end{aligned}$$

where $|\langle\psi_i|B|\phi_i\rangle| \le \|B|\phi_i\rangle\|$ since $|\psi_i\rangle$ is a unit vector and $\|B|\phi_i\rangle\| \le \|B\|$ follows from the definition of the operator norm. □

3 The Yao-Kremer Decomposition

A main tool in our analysis of quantum communication protocols is a theorem by Yao and Kremer. It essentially says that the state at the end of a quantum protocol with little communication has a 'simple' structure. For instance, if no communication takes place at all, the state of the system is clearly in a tensor state. If only a few qubits are transferred, the system is likely no longer in a tensor state, but can be written as the linear combination of a few tensor states.

THEOREM 4 *The final state of a communication protocol with q qubits of communication can be written as*

$$\sum_{m\in\{0,1\}^q} |\alpha_{x,m}\rangle|m_q\rangle|\beta_{y,m}\rangle,$$

where $|\alpha_{x,m}\rangle$ and $|\beta_{y,m}\rangle$ are vectors of norm at most 1 and $|m_q\rangle$ denotes the last bit of m (either 0 or 1).

PROOF: The proof is by induction on q. For $q=0$, the claim is obvious since the initial state of the system is a tensor product state. So assume the statement is true for q and let us prove it for $q+1$. Assume it is Alice's turn (the proof for Bob's turn is similar). Then the state after Alice applies U_{q+1} is

$$\sum_{m\in\{0,1\}^q} U_{q+1}(|\alpha_{x,m}\rangle|m_q\rangle)|\beta_{y,m}\rangle.$$

Define $|\alpha_{x,m0}\rangle = \mathrm{Tr}_{\mathcal{C}}\Pi_0 U_{q+1}(|\alpha_{x,m}\rangle|m_q\rangle)$, where Π_0 is the projection on $|0\rangle$ of the communication qubit, and $\mathrm{Tr}_{\mathcal{C}}$ indicates that we trace out the communication qubit (which is $|0\rangle$). We similarly define $|\alpha_{x,m1}\rangle = \mathrm{Tr}_{\mathcal{C}}\Pi_1 U_{q+1}(|\alpha_{x,m}\rangle|m_q\rangle)$, where Π_1 is the projection on $|1\rangle$ of the communication qubit. Then by definition, $|\alpha_{x,m0}\rangle|0\rangle + |\alpha_{x,m1}\rangle|1\rangle = U_{q+1}(|\alpha_{x,m}\rangle|m_q\rangle)$. Hence, by defining $|\beta_{y,m0}\rangle = |\beta_{y,m1}\rangle = |\beta_{y,m}\rangle$, we can write the state as

$$\sum_{m\in\{0,1\}^{q+1}} |\alpha_{x,m}\rangle|m_{q+1}\rangle|\beta_{y,m}\rangle,$$

as required. □

COROLLARY 5 *For any q-qubit protocol there exist numbers $a_{x,k}$, for $1\le k\le 2^{2q-2}$, $x\in X$, and similarly $b_{y,k}$ for $1\le k\le 2^{2q-2}$, $y\in Y$, with $|a_{x,k}|\le 1$,$|b_{x,k}|\le 1$ such that the acceptance probability of the input (x,y) is given by*

$$\sum_{k=1}^{2^{2q-2}} a_{x,k}\cdot b_{y,k}.$$

PROOF: By the previous theorem, we can write the acceptance probability as

$$\left(\sum_{\substack{m\in\{0,1\}^q\\ m_q=1}} \langle\alpha_{x,m}|\langle\beta_{y,m}|\right)\left(\sum_{\substack{m'\in\{0,1\}^q\\ m'_q=1}} |\alpha_{x,m'}\rangle|\beta_{y,m'}\rangle\right) = \sum_{\substack{m,m'\in\{0,1\}^q\\ m_q=m'_q=1}} \langle\alpha_{x,m}|\alpha_{x,m'}\rangle\langle\beta_{y,m}|\beta_{y,m'}\rangle$$

which is a sum over 2^{2q-2} terms. The claim follows. □

COROLLARY 6 *Let P be the $|X|\times|Y|$ matrix of acceptance probabilities of a q-qubit communication protocol. Then $\|P\|_{\rm tr} \leq 2^{2q-2}\sqrt{|X|\cdot|Y|}$, and hence*

$$q \geq \frac{1}{2}\log\frac{\|P\|_{\rm tr}}{\sqrt{|X|\cdot|Y|}}.$$

PROOF: Define an $|X|$-dimensional vector $|a_k\rangle = (a_{1k}, a_{2k}, \ldots, a_{|X|k})$ and a $|Y|$-dimensional vector $|b_k\rangle = (b^*_{1k}, b^*_{2k}, \ldots, b^*_{|Y|k})$. Then by the previous corollary,

$$P = \sum_{k=1}^{2^{2q-2}} |a_k\rangle\langle b_k|.$$

Hence,

$$\|P\|_{\rm tr} \leq \sum_{k=1}^{2^{2q-2}} \big\||a_k\rangle\langle b_k|\big\|_{\rm tr} = \sum_{k=1}^{2^{2q-2}} \big\||a_k\rangle\big\|\big\||b_k\rangle\big\| \leq 2^{2q-2}\sqrt{|X|\cdot|Y|}$$

where we used that the coordinates of $|a_k\rangle$,$|b_k\rangle$ are at most 1 in absolute value. □

4 Lower Bound on Inner Product

We now use the technique developed in the previous section to prove a lower bound on the quantum communication complexity of IP. Recall that in this problem $X = Y = \{0,1\}^n$ and we shall denote $N := |X| = |Y| = 2^n$. Define

$$M = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{\otimes n}$$

the matrix of inner products in ± 1 notation (instead of $\{0,1\}$). For example, for $n = 2$ the matrix is

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$

so, say, the bottom-right entry corresponds to $\mathsf{IP}(11, 11) = 0$. An important and easy to prove property that we shall use later is that all $N = 2^n$ singular values of M are equal to $2^{n/2}$. From the following lemma and Corollary 6 it follows that

$$Q(\mathsf{IP}) \geq \frac{1}{2}\log\frac{\|P\|_{\rm tr}}{\sqrt{2^n\cdot 2^n}} = n/4 - O(1).$$

LEMMA 7 *Any $N \times N$ matrix P whose entries are greater than $2/3$ whenever $\mathsf{IP}(x,y) = 0$ and less than $1/3$ otherwise, has trace norm $\|P\|_{\mathrm{tr}} \geq 2^{1.5n-1}/3$.*

PROOF: Consider the inner product $\langle P, M \rangle = \sum_{x,y} P_{xy}\, M_{xy}$. On the one hand, we know that $P_{xy} \geq 2/3$ whenever $M_{xy} = 1$ and $P_{xy} \leq 1/3$ whenever $M_{xy} = -1$. So $\langle P, M \rangle \geq 2^{2n-1}/3$. On the other hand, we have shown that $\langle P, M \rangle \leq \|P\|_{\mathrm{tr}} \cdot \|M\| = 2^{n/2}\|P\|_{\mathrm{tr}}$. Therefore, $\|P\|_{\mathrm{tr}} \geq 2^{1.5n-1}/3$. $\square$

We remark that the best known lower bound is roughly n, and that it is essentially tight by the trivial protocol where Alice sends her entire input to Bob (see the homework for one result in this direction). In the model where shared entanglement between the parties is allowed, the best known lower bound is roughly $n/2$, and is again essentially tight because of super-dense coding.

5 Razborov's Lower Bound on Disjointness

A crucial component in the proof of the previous section was that the singular values of M are all very small. This, together with the fact that $\langle P, M \rangle$ is large, allowed us to conclude that $\|P\|_{\mathrm{tr}}$ must be large. However, in the matrix M corresponding to DISJ, the singular values are no longer all very small. Our proof below is therefore slightly more sophisticated. In particular, instead of considering just one inner product $\langle P, M \rangle$, we will consider several inner products $\langle P, \mu_0 \rangle, \langle P, \mu_1 \rangle, \ldots, \langle P, \mu_{n/8} \rangle$.

Let $X = Y \subseteq \{0,1\}^n$ contain all strings of Hamming weight $n/4$. So $N := |X| = |Y| = \binom{n}{n/4}$. For $s = 0, 1, \ldots, n/4$, let μ_s be the $N \times N$ (symmetric) matrix corresponding to the uniform distribution over $\{(x,y) \in X \times Y \ : \ |x \cap y| = s\}$. In other words, the (x,y) entry in μ_s is nonzero iff $|x \cap y| = s$, and all nonzero entries are equal to the reciprocal of

$$\binom{n}{n/4}\binom{n/4}{s}\binom{n - n/4}{n/4 - s}$$

(which is the number of nonzero entries). Now assume that there exists a quantum communication protocol for DISJ with q qubits of communication, and let P be its acceptance probability matrix on inputs from $X \times Y$. The inner product $\langle P, \mu_s \rangle$ gives the average acceptance probability of the protocol on input two random sets with intersection s. Hence, by the correctness of the protocol, we must have that $\langle P, \mu_0 \rangle \in [2/3, 1]$ and that $\langle P, \mu_s \rangle \in [0, 1/3]$ for $s = 1, \ldots, n/4$. The main part of the proof is in proving the following lemma.

LEMMA 8 *For every $d \leq n/4$ there exists a degree d polynomial p such that for all $0 \leq s \leq n/8$,*

$$\big|p(s) - \langle P, \mu_s \rangle\big| \leq \frac{\|P\|_{\mathrm{tr}}}{N \cdot 2^{d/4}}.$$

Let us see how to complete the proof using this lemma. Choose $d = 8q + 100$. By Corollary 6, the right hand side of the above inequality is at most $2^{2q}N/(N2^{d/4}) \leq 0.01$. Therefore, we obtain a degree d polynomial satisfying that $p(0) \in [-0.01, 0.35]$ and that $p(s) \in [0.65, 1.01]$ for $s = 1, \ldots, n/8$. A theorem by Paturi (which is also used in the lower bound on Grover's problem) says that such a polynomial must be of degree $\Omega(\sqrt{n})$; so we obtain $q = \Omega(\sqrt{n})$ as required. It remains to prove the lemma.

PROOF: The matrices μ_s are known as combinatorial matrices, and it is known that they all share the same eigenspaces $E_0, E_1, \ldots, E_{n/4}$ (we will not prove this fact here). Hence there exists an orthogonal matrix U

that simultaneously diagonalizes all the μ_s. In other words, for $s = 0, \dots, n/4$ we have

$$U\mu_s U^T = \begin{pmatrix} \lambda_{s,0} &&&&&&&&& \\ & \ddots &&&&&&&& \\ && \lambda_{s,0} &&&&&&& \\ &&& \lambda_{s,1} &&&&&& \\ &&&& \ddots &&&&& \\ &&&&& \lambda_{s,1} &&&& \\ &&&&&& \ddots &&& \\ &&&&&&& \lambda_{s,n/4} && \\ &&&&&&&& \ddots & \\ &&&&&&&&& \lambda_{s,n/4} \end{pmatrix}$$

where the blocks correspond to the subspaces $E_0, \dots, E_{n/4}$. Another known fact is that $\lambda_{s,t}$ (which is the eigenvalue of μ_s in the subspace E_t) is a degree t polynomial in s and that

$$|\lambda_{s,t}| \le \frac{1}{N \cdot 2^{t/4}}$$

for $s \le n/8$. See Figures 2 and 3 for an example.

0.00006450	0.00006450	0.00006450	0.00006450	0.00006450	0.00006450
-0.00002150	-0.00000430	0.00001290	0.00003010	0.00004730	0.00006450
0.00000614	-0.00000319	-0.00000203	0.00000964	0.00003182	0.00006450
-0.00000142	0.00000164	-0.00000180	-0.00000018	0.00001806	0.00006450
0.00000024	-0.00000041	0.00000088	-0.00000233	0.00000602	0.00006450
-0.00000002	0.00000005	-0.00000014	0.00000061	-0.00000430	0.00006450

Figure 2: The values $\lambda_{s,t}$ for $n = 20$, with rows corresponding to $t = 0, \dots, n/4$ and columns to $s = 0, \dots, n/4$. Notice the fast decay in the left columns.

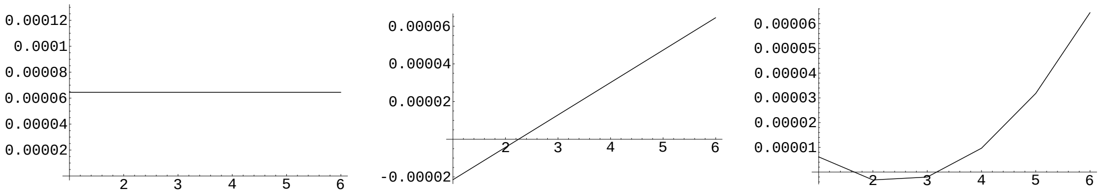

Figure 3: Plots showing the first three rows of the matrix from Figure 2.

Let us now consider P in the same basis. The resulting matrix UPU^T need not be diagonal. Still, for $t = 0, \dots, n/4$, let us define a_t as the trace of the diagonal block in UPU^T corresponding to the subspace E_t. Pictorially, it can be described as follows:

$$\begin{pmatrix} \boxed{a_0 \leftarrow \mathrm{tr}} &&& \\ & \boxed{a_1 \leftarrow \mathrm{tr}} && \\ && \ddots & \\ &&& \boxed{a_{n/4} \leftarrow \mathrm{tr}} \end{pmatrix}.$$

By definition, we have

$$\langle P, \mu_s \rangle = \langle UPU^T, U\mu_s U^T \rangle = \sum_{t=0}^{n/4} \lambda_{s,t} \cdot a_t$$

which is a degree $n/4$ polynomial in s. We now define a degree d polynomial by 'chopping' the above polynomial,

$$p(s) := \sum_{t=0}^{d} \lambda_{s,t} \cdot a_t.$$

Since the $\lambda_{s,t}$ decay exponentially with t (as long as s is not too big), we can show that p is a good approximation to the original polynomial. More precisely, for $s = 0, \ldots, n/8$,

$$|p(s) - \langle P, \mu_s \rangle| = \Big| \sum_{t=d+1}^{n/4} \lambda_{s,t} \cdot a_t \Big| \leq \frac{1}{N \cdot 2^{d/4}} \sum_{t=d+1}^{n/4} |a_t| \leq \frac{1}{N \cdot 2^{d/4}} \sum_{t=0}^{n/4} |a_t|.$$

The sum $\sum_{t=0}^{n/4} |a_t|$ can be seen as the inner product of UPU^T with a diagonal matrix with ± 1 on the diagonal. By Lemma 3, this sum is at most $\|P\|_{\mathrm{tr}}$, and the proof is complete. $\square$

Acknowledgement

I thank Ronald de Wolf for many comments and helpful tips.

References

[1] S. Aaronson and A. Ambainis. Quantum search of spatial regions. *Theory of Computing*, 1(4):47–79, 2005.

[2] H. Buhrman, R. Cleve, and A. Wigderson. Quantum vs. classical communication and computation. In *Proc. 30th ACM Symp. on Theory of Computing (STOC)*, pages 63–68, 1998.

[3] B. Kalyanasundaram and G. Schnitger. The probabilistic communication complexity of set intersection. *SIAM J. Discrete Math*, 5(4):545–557, 1992.

[4] A. Razborov. Quantum communication complexity of symmetric predicates. *Izvestiya of the Russian Academy of Science, mathematics*, 67(1):159–176, 2003. quant-ph/0204025.

[5] A. C.-C. Yao. Some complexity questions related to distributive computing. In *Proceedings of 11th ACM STOC*, pages 209–213, 1979.

[6] A. C.-C. Yao. Quantum circuit complexity. In *Proceedings of 34th IEEE FOCS*, pages 352–360, 1993.